



**University of
Zurich**^{UZH}

**Zurich Open Repository and
Archive**

University of Zurich
University Library
Strickhofstrasse 39
CH-8057 Zurich
www.zora.uzh.ch

Year: 2018

Comments on: Some recent work on multivariate Gaussian Markov random fields

Sain, Stephan R ; Furrer, Reinhard

DOI: <https://doi.org/10.1007/s11749-018-0609-z>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-160543>

Journal Article

Accepted Version

Originally published at:

Sain, Stephan R; Furrer, Reinhard (2018). Comments on: Some recent work on multivariate Gaussian Markov random fields. *Test*, 27(3):545-548.

DOI: <https://doi.org/10.1007/s11749-018-0609-z>

Comments on: Some recent work on multivariate Gaussian Markov random fields

Sain, Stephan R.¹ and Furrer, Reinhard²

In *Some recent work on multivariate Gaussian Markov random fields*, author Ying C. MacNab unifies several lines of research focused on multivariate formulations of Gaussian Markov random field (MRF) models through a coregionalization framework. MRFs are natural models for data on regular or irregular lattices, and, as the author has noted, they have found application in a wide range of scientific areas. There are significant computational advantages to MRFs that allow consideration of much larger datasets, but there are a number of other challenges that arise with such models. The author specifically addresses two of these, namely the entanglement of spatial and non-spatial components and the enforcement for positivity condition.

Multivariate Gaussian MRFs involve the specification of a precision matrix that encompasses both the dependence between variables and the spatial dependence across location. This precision matrix is often sparse, which enables the use of sparse matrix methods to improve computational performance. However, these precision matrices have very complex structure and parameterizations, which leads to the noted issues with entanglement of the spatial and non-spatial components and difficulty with ensuring these matrices are positive definite. While the paper yields a great deal of information about these issues, there are some practical concerns for those interesting in using such models that are also important. We comment on those in the following section. Next, against a statement made in the paper, the model of Sain et al. (2011) does allow a separable structure as we show. Finally, we conclude with a further gap that statisticians should close in order to further unify MRFs.

Practical issues

Much of the paper involves unifying the different types of multivariate MRFs through coregionalization, and a number of different types are presented. This is an impressive undertaking and yields some interesting insights with respect entanglement of the spatial and non-spatial dependence.

¹Jupiter, Boulder, USA

²Department of Mathematics, University of Zurich, Zurich, Switzerland
Department of Computational Science, University of Zurich, Zurich, Switzerland

However, it seems that there is still not one unified model formulation that allows movement between the different approaches through some set of parameters. Thus an analyst is stuck with choosing one model and will likely not try different alternatives. Or, the analyst is tasked with implementing a number of different formulations and then trying to choose objectively in a manner similar to what was done in the examples in the paper. It is still important to provide guidance to analysts with the practical implications of the different models and, in particular, encourage the analyst to seek out different options. Of course, it remains an active research area to develop objective procedures to choose between competing models.

One important reason to consider using MRFs is the ability to capitalize on sparse matrix methods, particularly for larger datasets. The two examples discussed are useful illustrations, but they have less than 100 locations each, which, for many applications, particularly in climate and weather, remote sensing, etc. is really quite small. If the ability to use sparse methods is lost, then it is a serious drawback to this approach. While Section 6 of the paper deals with some of the computational issues, it seems that the focus was on enforcing positivity of the precision matrix. While computation for fairly small sample sizes such as those in the examples does not really need sparse methods, computational approaches that are able to handle larger datasets are critically important.

Another key issue with MRF models is that they do not, in general, lead to stationary models. More importantly, this is often due to the configuration of the lattice, which, is often completely unrelated to the underlying physical process. Further, different model formulations lead to different nonstationarities. Sain et al. (2010) used a kronecker product of two univariate stationary MRFs for a regular lattice. While stationary and very computationally efficient for larger sample sizes, the resulting covariance function has some less than desirable properties. It would be interesting to see if these different types of coregionalization models are stationary and how they compare with each other.

The paper presents only first-order neighborhood models. For smoother fields or when approximating Gaussian random fields with a Gaussian MRF, higher-order neighbor-structures are needed (Rue and Tjelmeland, 2002). Such neighborhood structures are conceptually similar for the different models presented in the paper but we presume that they lead to different practical implications. Different neighborhood structures add another level of complexity to what are already fairly complex statistical models. However, it would be useful to see whether such models fit into a ready-to-use Bayesian software environment. Further, a discussion of likelihood estimation would

also be useful to help address issues such as choice of parameterization or the impact or transformations or constraints for parameters to improve estimation.

Separability

The author notes in Section 3.2 that, “The Sain et al. (2011) framework does not contain separable models in its family of MGMRFs.” This is not generally the case. Consider the bivariate example outlined in the Sain et al. (2011) paper with dependency connections defined to capture: 1) dependence between variables at a particular location, 2) spatial dependence within a variable, and 2) cross-spatial dependence. After some simplification, Eqn (5) of Sain et al. (2011) is written as

$$\left[I_n \otimes \tau^{1/2} \right] \left[I_n \otimes A - \begin{bmatrix} 0 & B\delta_{ij} \\ & \ddots \\ B'\delta_{ij} & 0 \end{bmatrix} \right]^{-1} \left[I \otimes \tau^{1/2} \right],$$

where $\tau = [\tau_1, \dots, \tau_p]'$ is a p -vector of scale parameters where p is the multivariate dimension, A is a $p \times p$ symmetric matrix with 1s on the diagonal and representing the “within-location” dependence structure, B is a $p \times p$ matrix of spatial dependence parameters, and δ_{ij} representing neighborhood structure (i.e., $\delta_{ij} = 1$ location i and location j are neighbors and $\delta_{ij} = 0$ otherwise). Further restricting B to be symmetric (and removing potential for asymmetric cross-dependence) and setting $B = \phi A$ where ϕ can now be thought of as a general spatial dependence parameter, Eqn (5) from Sain et al. (2011) can be simplified to

$$\left[I_n \otimes \tau^{1/2} \right] \left[I_n \otimes A - \phi \begin{bmatrix} 0 & A\delta_{ij} \\ & \ddots \\ A\delta_{ij} & 0 \end{bmatrix} \right]^{-1} \left[I \otimes \tau^{1/2} \right].$$

With some rearranging, this becomes $\left[I_n \otimes \tau^{1/2} \right] [(I_n - \phi H) \otimes A]^{-1} \left[I \otimes \tau^{1/2} \right]$, which gives the necessary Kronecker form between the spatial component $((I_n - \phi H))$ and the non-spatial form (A) .

The Sain et al. (2011) framework requires the specification of connections both within and across the lattices associated with the different variables. It also requires different assumptions and constraints on what values the parameters associated with those connections can take on. Together these two steps are ultimately controlled by the modeler or analyst and specify the model structure, and it is up to the modeler or analyst to carefully consider

these two steps to span the different model types. There are different model structures under this formulation that would have a separable form (such as the example discussed in Sain et al. (2011) and its extension to more than two variables), although it is likely that this is not true in general. There are lattice structures and different specifications of connections (or no connections, which is equivalent to forcing certain parameters to have a zero value) that can easily be seen to not have a separable form.

References

Rue, H. and Tjelmeland, H. (2002), “Fitting Gaussian Markov Random Fields to Gaussian Field,” *Scandinavian Journal of Statistics*, 29, 31–49.

Sain, S. R. and Furrer, R. (2010), “Combining Climate Model Output via Model Correlations,” *Stochastic Environmental Research and Risk Assessment*, 24(6), 821–829.

Sain, S. R., Furrer, R., Cressie, N. (2011), “A Spatial Analysis of Multivariate Output from Regional Climate Models,” *Annals of Applied Statistics*, 5(1), 150–175.

A conditional univariate approach

Reinhard? This is Noel's comment, right? Can you add anything here? Or should we just forget it since I think it is just you and I now anyway.

There is an alternative conditional approach, and some of the work the three of us have done uses such an approach. The paper by Cressie and Zammit-Mangion (2016, *Biometrika*) makes the case for a conditional approach on *whole geostatistical processes*. There, each part of the model is conditionally univariate, with spatial dependence captured by a spatial covariance function in practice this is a covariance matrix on a fine-resolution grid. It would be simple to change the fine-resolution grid to a collection of small areas and the series of conditional covariance matrices to a series of conditional (sparse) precision matrices over the small areas. If that is done, I think one obtains some of the models that the three of us worked on.

Old Stuff

This short notes deals with the assertion that the model in the Sain et al. (2011) paper does allow a separable form. By this, we mean that the covariance matrix can be written as a Kronecker form:

$$\Sigma = \text{Space} \otimes \text{Multivariate}.$$

Eqn (5) from Sain et al. (2011):

$$\left[I_n \otimes \tau^{1/2} \right] \left[I_n \otimes A - \begin{bmatrix} 0 & & B\delta_{ij} \\ & \ddots & \\ B'\delta_{ij} & & 0 \end{bmatrix} \right]^{-1} \left[I \otimes \tau^{1/2} \right].$$

Since the scale parameters in τ are both pre- and post-multiplied, they do not impact this discussion. We will focus on the A and B matrices. First, let B be symmetric, i.e. $B = B'$. The second step would be to restrict the within-variable spatial parameters ϕ_{ii} to be the same, i.e.

$$-\phi_{11} = \dots = -\phi_{pp} = \phi.$$

Now the center of Eqn (5) can be simplified to

$$I_n \otimes A - \phi H \otimes B,$$

where H is a neighbor matrix with 0s except where location i is a neighbor of location j in which case $H_{ij} = 1$. The matrix B now has the form

$$B = \begin{bmatrix} 1 & & -\phi_{jl}/\phi \\ & \ddots & \\ -\phi_{jl}/\phi & & 1 \end{bmatrix}.$$

The matrix B has the same dimension as A , and if we equate the off-diagonal elements, i.e. $\rho_{jl} = \phi_{jl}/\phi$, then we have $A = B$. This implies that the center of Eqn (5) can be written as $(I_n - \phi H) \otimes A$, which gives us the Kronecker form of a separable covariance.

There are two things to note here. First, the Sain et al. (2011) formulation gives a rich class of models and, depending on how you specify connections both within and across the lattices associated with each variable and depending on how you constrain the parameters that are associated with those connections, a separable form is possible. Second, while we may be able to demonstrate separability in this case, it is not clear that it is true in general. It is not hard to imagine cases where the connections in the lattices are defined or omitted (equivalent to setting certain parameters to zero) in such a way to prevent this case.